基于超像素统计量的随机森林遥感图像分类 *

石彩霞 [1]，赵传钢 [1]，庞　蕾 [2]

(1. 北京林业大学 信息学院，北京 100083; 2. 北京建筑大学 测绘与城市空间信息学院，北京 100044)

摘　要：针对遥感图像地物覆盖分类方法对图像空间分布信息利用不足的问题，提出一种基于超像素统计量的随机森林遥感图像分类方法。以北京市海淀区为研究区，选用 Landsat-8 卫星为主要数据源，通过改进 SLIC 超像素分割方法，使之适用于多光谱遥感图像中超像素的分割，提取超像素常见的六个统计量(最小值、最大值、均值、标准差、上四分位数、下四分位数)用于随机森林在遥感图像中的分类。实验结果表明，本文对研究区遥感图像的总体分类精度为 89.01%，明显改善了对地物的错分和漏分现象，能够推广到 Landsat-8 遥感图像的地物覆盖分类工作中。

关键词：Landsat-8；随机森林；超像素；地物覆盖；简单线性迭代聚类

中图分类号：TP75　　doi: 10.3969/j.issn.1001-3695.2017.07.0677

Research on random forest remote sensing image classification based on superpixel statistics

Shi Caixia[1], Zhao Chuangang[1], Pang Lei[2]

(1. *School of Information Science & Technology*, *Beijing Forestry University*, *Beijing* 100083, *China*; 2. *School of Geometrics & Urban Information*, *Beijing University of Civil Engineering & Architecture*, *Beijing* 100044, *China*)

Abstract: In order to solve the problem of insufficient utilization of the spatial distribution information in remote sensing image classification, this paper proposed a random forest remote sensing image classification method based on superpixel statistics. Taking Haidian district, Beijing as study area and Landsat-8 OLI (Operational Land Imager) image as the main data source, the author utilized six statistics(maximum, minimum, mean, standard deviation, upper quartile, lower quartile) calculated by superpixel segmentation algorithm as features to build remote sensing image classification model based on Random Forest. The experiments show that superpixel statistical features can significantly improve the overall classification accuracy of remote sensing image to 89.01%, and effectively reduce omission error and commission error. This method can be extended to land cover classification from Landsat-8 remote sensing image.

Key Words: Landsat-8; random forest; superpixel; land cover; simple linear iterative cluster (SLIC)

0　引言

遥感信息提取的关键技术之一就是遥感图像分类。过去的几十年里，国内外专家和学者一直致力于提高遥感图像分类精度。遥感图像分类主要分为基于像素的分类和基于对象的分类。基于像素的分类方法以像素为基本处理单元，忽视了图像中邻近像素间的空间分布关系，限制了分类精度的提高。基于对象的分类方法充分利用了图像的光谱信息和空间分布信息，其分类性能优于基于像素的分类方法[1, 2]。

图像分割是基于对象分类方法的关键步骤之一，分割质量的好坏直接影响分类精度[3]。超像素是近年来快速发展的一种图像分割方法，相较于以往的图像分割方法，更有利于图像局部特征的提取，而且能有效降低图像后续处理的计算复杂度[4, 5]。超像素分割方法分为两类，一类是基于图论的方法：NC（normalized cuts）[6]、GS（graph based）[7]、SL（superpixel lattice）[8]；另一类是基于梯度下降的方法：WS（watersheds）[9]、MS（Mean-Shift）[10]、QS（quick shift）[11]、TP（Turbopixels）[12]、SLIC（simple linear iterative clustering）[13]等。其中 SLIC 算法相比于其他算法，具有分割速度快、计算复杂度低、超像素边缘贴合度和紧密度较好、生成超像素数量可控制的优点。现已有研究将 SLIC 算法应用于遥感图像的分割处理[14, 15]，但大多是针对彩色遥感图像的分割，而鲜见于多光谱遥感图像的分割处理。

随机森林（random forest，RF）算法作为一种基于分类与回归决策树的机器学习算法[16]，在遥感图像分类领域受到国内外研究学者的广泛关注[17-19]。特征变量的选择对 RF 算法的分

基金项目：建筑学基金资助项目（61501019）；北京市教育委员会科技计划一般项目（SQKM201610016008）

作者简介：石彩霞（1991-），女，山西吕梁人，硕士研究生，主要研究方向为计算机应用、图像处理、机器学习（scx_fighting@163.com）；赵传钢（1971-），男，副教授，博士，主要研究方向为信息检索、机器学习等；庞蕾（1971-），女，副教授，博士，主要研究方向为多基线干涉 SAR 数据处理、层析 SAR 遥感技术等.

类精度有重要影响。文献[20]将光谱、形状和纹理特征共同作为 RF 算法的特征变量应用于高分辨率遥感图像，有效改善了总体分类精度。文献[21]将 Getis 统计量作为特征变量加入 RF 分类模型中提高了 RF 算法的分类性能。文献[22]提出将矩距离指数（Moment Distance Index，MDI）、光谱特征、地形特征相结合用于帕米尔高原的植被覆盖分类。目前 RF 算法中特征变量的选择没有较通用的方法用于各种类型的遥感图像，因此 RF 分类算法在特征变量的选择上仍有很大的改善空间。

本文改进了 SLIC 超像素分割算法使之适用于多光谱遥感图像，提取超像素常见的 6 个统计量用于 RF 算法在多光谱遥感图像的地物分类中，从而有效利用遥感图像的光谱信息和空间信息。

1　相关方法及原理

1.1　随机森林

RF 是由决策树分类器 $\{h(\mathbf{x},\Theta_k),k=1,\mathbf{K}\}$ 组合的集成分类器，最终的分类结果是由基分类器投票决定。RF 算法的基本流程是：

a)用 Bootstrap 抽样方法从原始样本集中随机抽取 K 个训练样本集。

b)个训练样本集分别构建决策树模型，得到 K 种分类结果。

c) K 种分类结果采用投票表决的方式确定最终分类结果。

RF 算法基本流程如图 1 所示。

步骤 b)每棵决策树的输入变量是从 N 个特征中随机抽取 M 个特征。通常 M 取值为 $M=\sqrt{N}$ 或 $M=\log_2 N+1$ （N 为总的特征数目）。另外，节点分裂标准通常采用基尼指数（Gini）和信息熵（Entropy）。两种方法的计算公式分别为：

$$Gini(D)=1-\sum_{i=1}^{m}p_i^2 \tag{1}$$

$$Entropy(D)=-\sum_{i=1}^{m}p_i\log_2(p_i) \tag{2}$$

其中：m 表示训练数据集 D 包含的类别个数，p_i 表示训练数据属于类别 C_i 的概率。

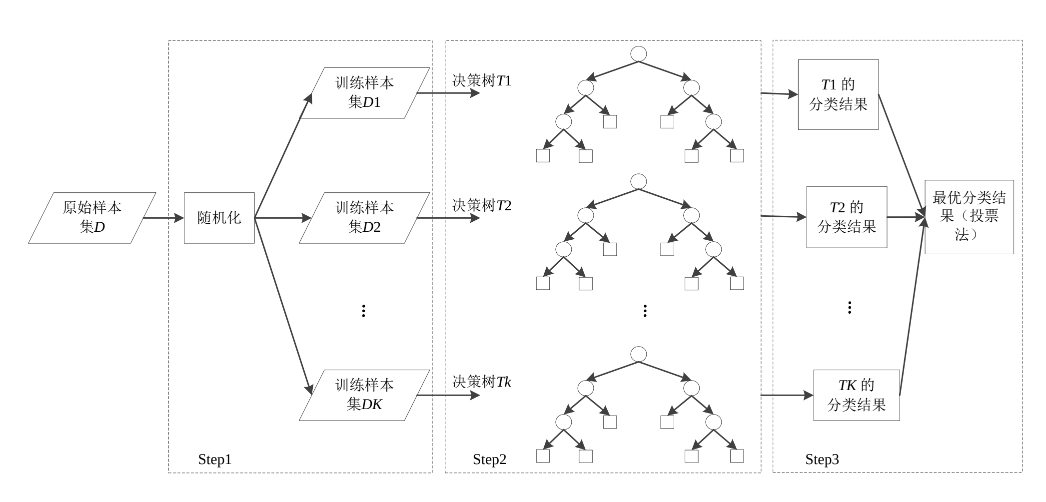


图 1　RF 算法基本流程

1.2　超像素分割

1.2.1 SLIC 超像素分割算法

SLIC 算法[13]是一种基于梯度下降的超像素分割方法，根据各像素之间空间和颜色距离的相似性进行分割，整个分割过程只需要一个参数 K 用于控制生成的超像素的数量。

SLIC 算法的基本流程如下：

a)初始化种子点。在像素点个数为 N 的图像上，按照步长 $S=\sqrt{N/K}$ 均匀分布种子点。在原始种子点的 $3*3$ 邻域内，将梯度值最小的像素点作为新的种子点，同时为每个种子点分配一个单独的标签值。

b)为每个像素点分配标签值。在以种子点为中心，$2S*2S$ 的搜索范围内，计算该种子点与搜索范围内各像素点的距离，然后将与各像素点距离最小的种子点的标签值赋给对应的像素点。

c)更新种子点。计算每个超像素块的坐标重心，将其作为超像素块新的种子点的位置。

d)重复步骤 b)c)，直至结果收敛。

e)去掉多连通以及过小超像素。

步骤 b)中，距离相似性的度量维度是 5 维（CIELAB 色彩空间：$[l\quad a\quad b]^{\mathrm{T}}$ 和坐标空间：$[x\quad y]^{\mathrm{T}}$），其计算公式如下：

$$d_c^{ij}=\sqrt{(l_j-l_i)^2+(a_j-a_i)^2+(b_j-b_i)^2} \tag{3}$$

$$d_{xy}^{ij}=\sqrt{(x_j-x_i)^2+(y_j-y_i)^2} \tag{4}$$

$$D^{ij}=\sqrt{\left(\frac{d_c^{ij}}{m}\right)^2+\left(\frac{d_{xy}^{ij}}{S}\right)^2} \tag{5}$$

其中：d_c^{ij} 为像素点间的颜色差异，表示像素 i 和像素 j 在色彩空间 $[l\quad a\quad b]^{\mathrm{T}}$ 的距离；d_{xy}^{ij} 为像素点间的空间距离；S 为种子点的间距；m 为常数，用来衡量颜色相似性和空间邻近性在距离计算中的相对重要性；D^{ij} 为两个像素的距离，即相似度，D 值越小，说明两个像素越相似。

1.2.2 改进的 SLIC 超像素分割算法

原 SLIC 算法是针对彩色图像的处理，而研究数据为多光谱遥感图像（7 个波段），将距离度量的维度由原先 5 维扩展为 9 维（色彩空间：波段 1-7：$[b_1\ \ b_2\ \ b_3\ \ b_4\ \ b_5\ \ b_6\ \ b_7]^{\mathrm{T}}$ 和坐标空间：$[x\ \ y]^{\mathrm{T}}$），由此得到距离的计算公式如下：

$$d_{band}^{ij}=\sqrt{\sum_{k=1}^{7}(b_{kj}-b_{ki})^2} \tag{6}$$

$$D^{ij}=\sqrt{\left(\frac{d_{band}^{ij}}{m}\right)^2+\left(\frac{d_{xy}^{ij}}{S}\right)^2} \tag{7}$$

其中：d_{band}^{ij} 为像素点间的颜色差异，表示像素 i 和像素 j 在波段空间的距离；b_{kj} 和 b_{ki} 分别表示像素 i 和像素 j 在波段 k 的像元值。同理，D^{ij} 表示两个像素的相似度。

1.2.3 超像素统计量

利用改进的 SLIC 超像素分割算法计算得到超像素，然后对遥感图像中超像素的特征进行提取，并将这些特征作为 RF 分类模型的分类特征向量，以达到充分利用遥感图像的空间分布信息的目的。

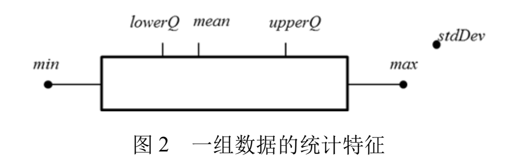


图 2　一组数据的统计特征

本文从统计分析的角度进行超像素的特征提取，根据图 2 所示的箱体图，一组数据的总体分布特征一般包括 6 个统计量：最小值（min）、最大值（max）、均值（mean）、上四分位数（upperQ）、下四分位数（lowerQ）和标准差（stdDev）。将这 6 个统计量共同作为特征来描述超像素的整体信息。

1.3　RF 与超像素统计量相结合的遥感图像分类

1.3.1 算法描述

本文将改进的 SLIC 超像素分割算法与 RF 算法相结合，把超像素的统计量特征作为分类特征输入到 RF 模型中进行遥感图像的地物分类。本文采用技术路线如图 3 所示。

其中，主要处理步骤为：

a）先对 Landsat8 OLI 多光谱遥感图像进行预处理，包括对图像进行辐射定标、大气校正、工程区图像的镶嵌与裁剪；

b)采用改进的 SLIC 超像素分割算法对遥感图像进行分割，将特征相似的在一定区域内的像素点聚类成超像素块；对每个超像素在每个波段上提取六个统计量（min，max，mean，stdDev，upperQ，lowerQ）作为超像素区域内各个像素的统计特征；

c)将多光谱波段数据、纹理特征与超像素统计特征相结合，共同作为输入变量构建 RF 分类模型，得到地物分类图；

d)将分类结果进行分类精度的验证。

步骤 b)中，利用改进的 SLIC 算法进行超像素分割，分割如图 4 所示。对图像中的每一个超像素块提取 7 个多光谱波段上的 7×6=42 个统计特征，提取示意图如图 5 所示。

步骤 c)中，纹理特征由 GLCM[23]计算得到，包括均值（mean）、方差（variance）、协同性（homogeneity）、对比度（contrast）、相异性（dissimilarity）、信息熵（entropy）、二阶矩（angular second moment）、相关性（correlation）共八个特征，每个纹理特征在 Landsat-8 的七个多光谱波段分别进行计算。

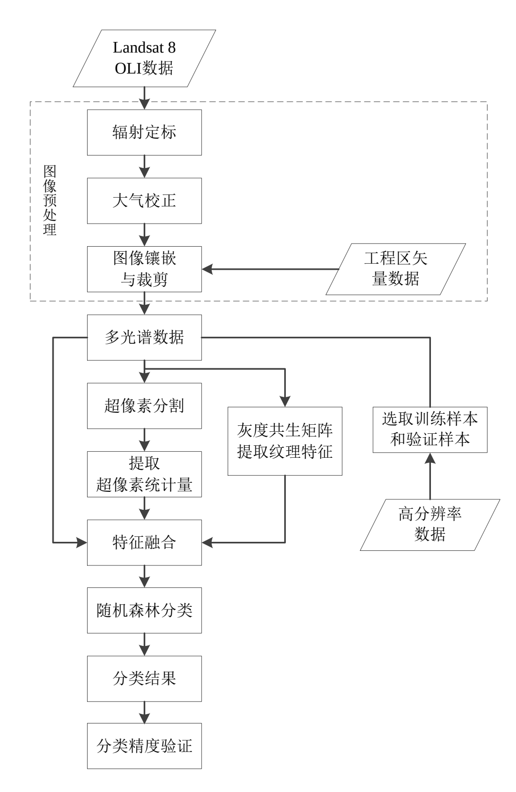


图 3　技术路线图



图 4　SLIC 超像素分割

1.3.2 算法复杂度分析

本文主要选择 SLIC 超像素分割和随机森林分类两个算法用于遥感图像分类。假设图像包含的像素点数为 N，随机森林算法中决策树个数为 K，分类特征数为 M，SLIC 分割算法的时间复杂度为 $\mathrm{O}(N)$ [13]，随机森林分类算法的时间复杂度近似为 $\mathrm{O}(KMN(\log N)^2)$ [24]。

2　实验

2.1　研究区概况、分类系统及数据源

2.1.1 研究区概况与分类系统

本文以北京市海淀区为研究区（北纬 39°53′~40°09′，东经 116°03′~116°23′）（图 6）。海淀区气候属温带湿润季风气候区，

冬季寒冷，夏季炎热多雨。全区面积为430.77平方千米，地势西高东低，西部为山地，东部和南部为平原。研究区西部海拔较高的地区覆盖类型是林地，农田与分散的居民点呈混合分布的状态，区域内有大小河流10条，还有玉渊潭、上庄水库等湖泊。研究区具有典型的城市地物类型特点，据此将本研究区作为实验区域。

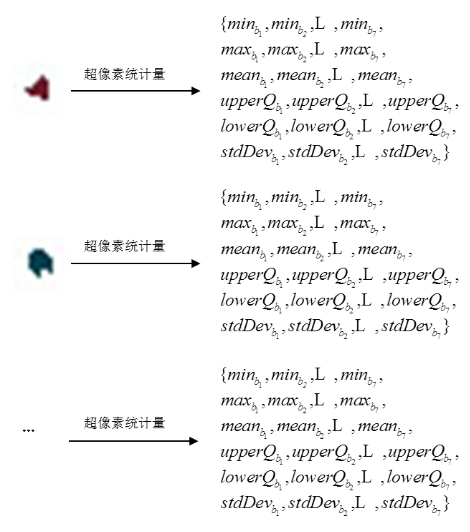


图5　超像素统计量分析

根据全国土地遥感监测利用/覆盖分类体系和遥感数据土地利用/覆被分类体系，结合北京市海淀区的土地利用现状以及本文所使用遥感数据源（Landsat-8）的特点，将研究区的地物覆盖类型分为水体、耕地、林地、建设用地和裸地五类。

2.1.2 数据源

本研究使用的遥感数据是从美国地质勘探局（United States Geological Survey, USGS）网站上下载的经过L1T级地形校正的Landsat-8卫星OLI多光谱影像。研究区需要一景Landsat-8 OLI数据覆盖，对原始影像（轨道号为123/032，成像时间为2013年10月3日）进行相应的图像预处理，以便进行后续的图像处理操作。

本研究选用Google Earth高分辨率遥感影像作为标注训练样本和验证分类精度的辅助数据。

2.2　实验过程

2.2.1 软件

本文使用ENVI5.3.1软件进行图像预处理操作，并将IDL作为开发语言，同时结合EnMAP-Box工具[25]实现了对本文研究的遥感图像分类方法的实验验证。实验平台采用2.10GHz处理器、6GB内存、Windows7 64位操作系统。

2.2.2 样本选择

结合目视解译（图7，解译图像为Landsat 8多光谱遥感图像的标准假彩色合成图），参考Google Earth提供的高分辨率影像选择训练样本和建模样本，同时保证样本在整个研究区域均匀分布，样本选择采用分层随机抽样的方法，具体情况见表1（注：训练样本和验证样本相互独立）。

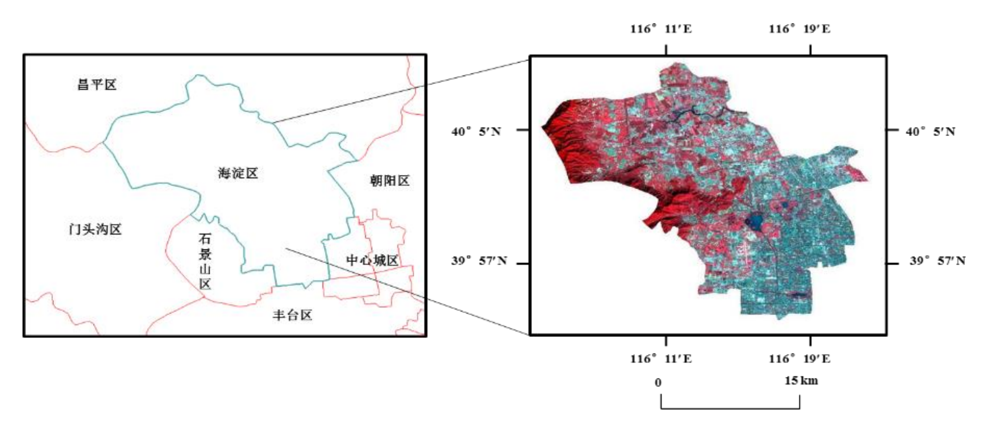
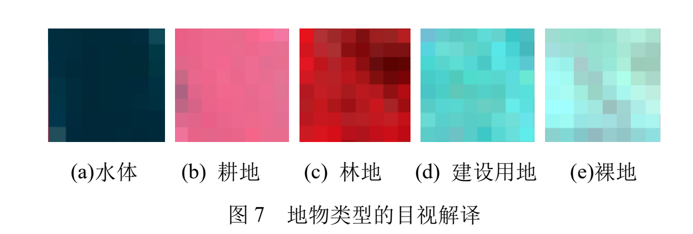


图6　研究区位置



(a)水体　(b) 耕地　(c) 林地　(d) 建设用地　(e)裸地

图7　地物类型的目视解译

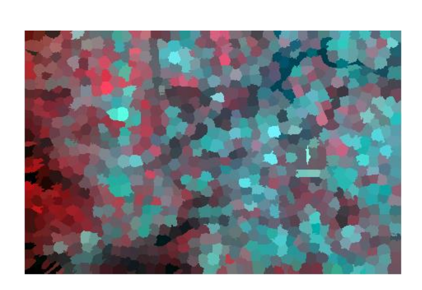
2.2.3 SLIC超像素分割及特征提取

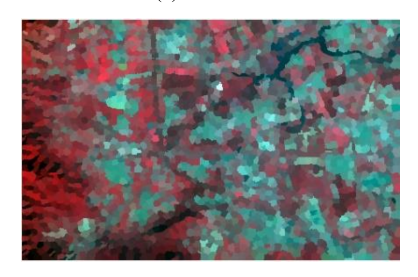
基于超像素提取图像的局部统计特征，分割尺度是影响统计特征提取的关键因素。假设超像素分割尺度为K，图像中共有N个像素点，预分割的每个超像素的大小为N/K。K值越大，生成的超像素块越小，相同的地物类型被聚类到同一区域，分割效果越好，时间需求递增，但超过一定的范围，可能会出现过分割，而且生成超像素的速度也明显降低。相反，K值越小，生成的超像素块越大，将不同的地物类型聚类到同一区域，分割效果越差，可能会导致欠分割。因此，选择合适的分割尺度需要在分割效果和分割速度之间取得平衡。

本文研究数据的空间分辨率为30m，研究区图像的大小为987*999像素，结合研究数据的特点以及图像的大小，分别使用K=10000，30000，50000，70000，90000的分割尺度提取超像素统计量。图8所示为研究区的局部区域在上述5种分割尺度下生成的超像素在波段5、4、3上提取均值特征得到的假彩色合成结果。

表 1　训练样本选择

类型	建模样本个数（像元）	验证样本个数（像元）
水体	100	286
耕地	725	562
林地	316	519
建设用地	1179	680
裸地	359	338
总计	2679	2385


(a) K=10000


(b) K=30000


(c) K=50000


(d) K=70000


(e) K=90000

图 8　不同分割尺度下的超像素分割结果（局部）

本文提取的超像素特征为各超像素在 7 个多光谱波段的 6 个统计量（min，max，mean，stdDev，upperQ，lowerQ），同时根据 GLCM 在 7 个多光谱波段上分别提取像素的 8 个纹理特征（纹理提取窗口为 5×5[22]），将超像素特征（7×6=42）、纹理特征（7×8=56）和 7 个多光谱波段数据赋予图像的各个像素构成每个像素的特征向量（共 105 个）。

为了进一步说明分割尺度对图像分类结果的影响，利用本文提出的分类方法对遥感图像在上述 5 种超像素分割尺度下进行分类，从而选择最合适的尺度提取超像素统计量，具体说明见 3.2 节。根据 3.2 节的实验结果，本文选择 K=50000 作为最终的超像素分割尺度。

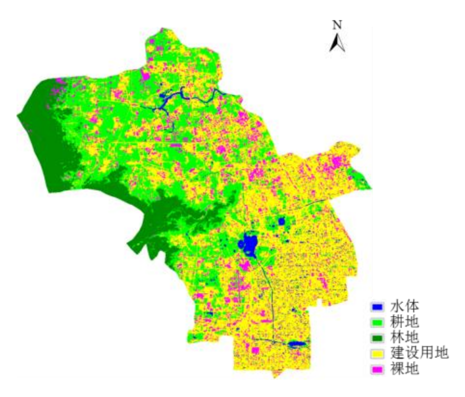
2.2.4 RF 分类模型构建

本文实验中，RF 分类模型的参数选择情况为：决策树的个数为 100，决策树特征子集的个数为 10（$\sqrt{105}$，本文候选特征数为 105），选择具有最小 Gini 指数的属性作为分裂属性。实验用由 7 个多光谱波段数据、SLIC 提取到的 42 个局部统计特征、GLCM 提取的 56 个纹理特征构造像素的特征向量（105 个），作为 RF 模型的输入变量。

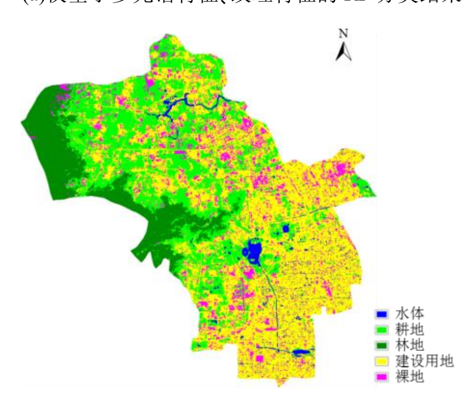
3　分类结果与精度验证分析

3.1　分类结果

图 9 所示为仅使用多光谱特征和纹理特征的 RF 分类结果和辅以超像素统计特征（分割尺度为 K=50000）的 RF 分类结果。



(a)仅基于多光谱特征、纹理特征的 RF 分类结果



(b)辅以超像素统计特征的 RF 分类结果(分割尺度 K=50000)

图 9　分类结果对比

3.2 精度验证

图 10 为不同分割尺度下提取的超像素统计特征用于分类得到的各地物类型的分类精度。其中，由于林地的超像素统计特征与其他地物差异性较大，5 种分割尺度对林地的提取精度基本一致。分割尺度对水体、耕地、建设用地和裸地的影响较大。K=50000 的分割尺度对水体、耕地、建设用地和裸地的提取精度较高。5 种分割尺度对水体、耕地、林地、建设用地和裸地的平均提取精度分别是 84.34%，87.01%，89.25%，85.47%，89.23%。

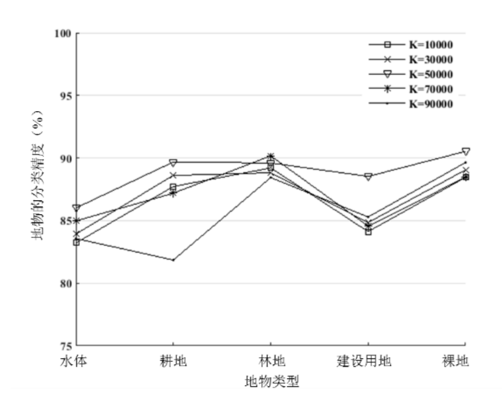


图 10　不同分割尺度下地物的分类精度

表 2 为仅使用多光谱特征和纹理特征以及辅以超像素统计特征（分割尺度 K=50000）的 RF 分类精度比较。与仅使用多光谱特征和纹理特征的 RF 算法相比，辅以超像素统计特征的 RF 算法将总体分类精度提高约 5.2%，Kappa 系数提高 7%。

表 2　分类精度比较

编号	分类方法	总体分类精度（%）	Kappa 系数
1	仅使用多光谱特征和纹理特征	83.77	0.79
2	辅以超像素统计特征（K=50000）	89.01	0.86

表 3 对仅使用多光谱特征和纹理特征的 RF 算法和辅以超像素统计特征（分割尺度为 K=50000）的 RF 算法的分类结果建立混淆矩阵后的精度评价。由表可知，辅以超像素统计特征的 RF 算法对水体、耕地、林地、建设用地和裸地这 5 类地物的分类精度与仅使用多光谱特征和纹理特征的 RF 算法相比，分别高约 5.94%，3.91%，1.55%，7.35%，8.28%。仅使用多光谱特征和纹理特征的 RF 算法对水体、裸地和建设用地的提取精度较低。其中，由于建设用地和裸地可能存在光谱混淆现象，造成两类地物一定程度的混分。而辅以 K=50000 分割的超像素统计特征，建设用地和裸地的错分率明显降低，主要是由于超像素统计特征能更大程度的将两类地物区分开；水体、林地和耕地的错分和漏分现象也得到一定程度的改善。超像素统计特征的加入对林地的提取精度提高不大。

3.3 变量重要性分析

特征变量重要性指特征变量对分类精度产生影响的程度。假设 RF 中有 N 棵决策树，对于变量 F，第 i 棵决策树 T_i 对应的 OOB（out-of-bag）数据计算得到 OOB 误差 OOB_{1i}；对 OOB 数据中所有样本的变量 F 随机加入噪声，计算 OOB 误差 OOB_{2i}，则变量 F 的重要性计算[25]采用公式：

$$Importance_F = \frac{\sum_{i=1}^{N} OOB_{2i} - OOB_{1i}}{N} \tag{8}$$

$Importance_F$ 越大，说明变量 F 的重要程度越大，在分类中起较为关键的作用。图 11 为由超像素提取的 42 个统计特征（分割尺度为 K=50000）和基于 GLCM 提取的 56 个纹理特征的变量重要性排序。这里仅取前 30 位进行展示。在重要性排名前 30 个特征变量中，有 24 个为基于超像素提取的统计特征，而基于 GLCM 的纹理特征仅为 6 个。排名前 3 位均为基于 GLCM 的均值特征，说明该特征的重要程度较大。虽然基于 GLCM 的均值特征排名靠前，但是基于 GLCM 的其他纹理特征重要性排名都处于较后位置，重要性远不及基于超像素提取的统计特征，对图像中不同地物的区分度不如超像素统计特征，因此本文提出的分类方法能提高分类精度，分类效果比仅使用多光谱特征和纹理特征的分类方法优良。

4　结束语

针对 Landsat-8 多光谱遥感图像提出一种基于超像素统计量的 RF 分类方法。本文对 SLIC 超像素分割方法进行扩展和改进，使之适用于多光谱遥感图像中超像素的生成。根据统计学的知识，对超像素提取 6 个统计量作为分类特征，充分利用了像素间的空间分布关系，有效提高特征辨识度。本文方法能有效提高算法的总体分类精度和 Kappa 系数，明显改善分类过程中对地物的错分和漏分现象，取得理想的分类结果。此外，变量重要性分析也充分表明超像素统计特征对分类精度的提高起到了重要的作用。

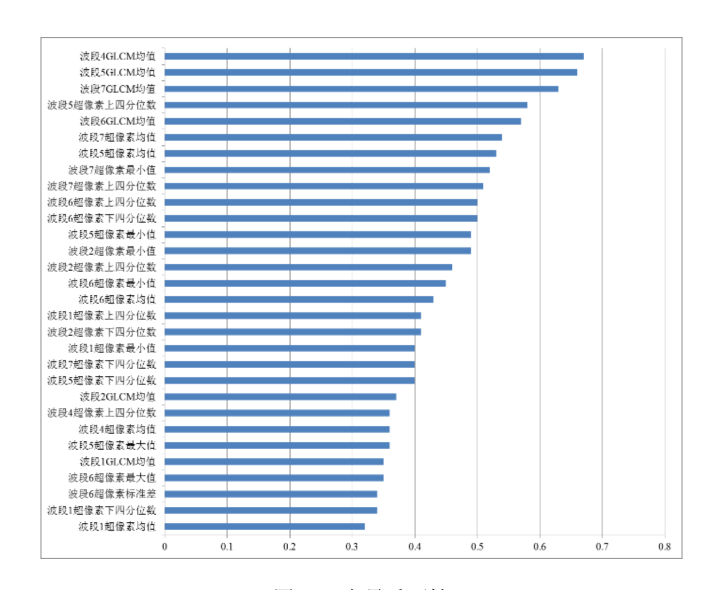


图 11　变量重要性

表 3　分类精度评价

类型	验证样本个数（像元）	分类正确样本数（像元）		分类精度（%）		错分率（%）		漏分率（%）	
		1	2	1	2	1	2	1	2
水体	286	229	246	80.07	86.01	4.98	2.77	19.93	13.99
耕地	562	482	504	85.77	89.68	17.75	12.8	14.23	10.32
林地	519	457	465	88.05	89.6	3.79	0.85	11.95	10.4
建设用地	680	552	602	81.18	88.53	20.69	13.88	18.82	11.47
裸地	338	278	306	82.25	90.53	28.17	20.73	17.75	9.47
总计	2385	1998	2123	83.77	89.01	-	-	-	-

注:表中编号 1 为仅使用多光谱特征和纹理特征的 RF 算法,编号 2 为辅以超像素统计特征(分割尺度 K=50000)的 RF 算法。

参考文献：

[1] Myint S, Gober P, Brazel A, et al. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery [J]. Remote Sensing of Environment, 2011, 115 (5): 1145-1161.

[2] Yu Q, Gong P, Clinton N, et al. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery [J]. Photogrammetric Engineering and Remote Sensing, 2006, 72 (7): 799-811.

[3] Liu D, Xia F. Assessing object-based classification: advantages and limitations [J]. Remote Sensing Letters, 2010, 1 (4): 187-194.

[4] 宋熙煜，周利莉，李中国，等. 图像分割中的超像素方法研究综述 [J]. 中国图象图形学报, 2015, 20 (5): 599-608.

[5] Ren Xiao-feng, Malik J. Learning a classification model for segmentation [C]// Procs of the 9th IEEE International Conference on Computer Vision. Washington DC: IEEE Computer Society, 2003: 10-17.

[6] Shi J, Malik J. Normalized cuts and image segmentation [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22 (8): 888-905.

[7] Felzenszwalb P, Huttenlocher D. Efficient graph-based image segmentation [J]. International Journal of Computer Vision, 2004, 59 (2): 167-181.

[8] Moore A, Prince S, Warrell J, et al. Superpixel lattices [C]// Proc of IEEE Conf. on Computer Vision and Pattern Recognition. 2008.

[9] Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1991, 13 (6): 583-598.

[10] Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2002, 24 (5): 603-619.

[11] Vedaldi A, Soatto S. Quick shift and kernel methods for mode seeking [C]. Proc. European Conf. Computer Vision, 2008.

[12] Levinshtein A, Stere A, Kutulakos K, et al. Turbopixels: fast superpixels using geometric flows [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2009, 31 (12): 2290-2297.

[13] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2012, 34 (11): 2274–2282.

[14] Yi Yuan, Xiangyu Hu. Random forest and objected-based classification for forest pest extraction from UAV aerial imagery [C]// Proc of ISPRS International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences, 2016: 1093-1098.

[15] 董志鹏，梅小明，陈杰，等. 简单线性迭代聚类的高分辨率遥感影像分割 [J]. 遥感信息, 2016, 31 (6): 83-88.

[16] Breiman L. Random forests [J]. Machine learning, 2001, 45 (1): 5-32.

[17] Lowe B, Kulkarni A. Multispectral image analysis using Random Forest [J]. International Journal on Soft Computing, 2015, 6 (1): 1-13.

[18] Rodriguez-Galiano V F, Ghimire B, Rogan J, et al. An assessment of the effectiveness of a random forest classifier for land-cover classification [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2012, 67 (1): 93-104.

[19] 刘毅，杜培军，郑辉，等. 基于随机森林的国产小卫星遥感影像分类研究 [J]. 测绘科学, 2012, 37 (4): 194-196.

[20] 刘海娟，张婷，侍昊，等. 基于 RF 模型的高分辨率遥感影像分类评价 [J]. 南京林业大学学报：自然科学版, 2015, 39 (1): 99-103.

[21] Ghimire B, Rogan J, Miller J. Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic [J]. Remote Sensing Letter, 2010, 1 (1): 45-54.

[22] Eric A L S, Kenneth G B, Raul V. Multispectral and texture feature application in image-object analysis of summer vegetation in eastern Tajikistan Pamirs [J]. Remote Sensing, 2016, 8 (78): 1-20.

[23] Haralick R M, Shanmugam K, Dinstein I. Textural features for image classification [J]. IEEE Trans on Man and Cybernetics, 1973, 3 (6): 610-621.

[24] 姚登举，杨静，詹晓娟. 基于随机森林的特征选择算法 [J]. 吉林大学学报：工学版, 2014, 44 (1): 137-141.

[25] Van Der Linden S, Rabe A, Held M, et al. The EnMAP-Box—A toolbox and application programming interface for EnMAP data processing [J]. Remote Sensing, 2015, 7 (9): 11249-11266.